Setting up Genesys Multicloud CX Private Edition

Architecture

2/11/2026

Contents

Understand the architecture and components of Genesys Multicloud CX private edition; the supported third-party back-end services; and how they all work together in both single- and multi-region deployments.

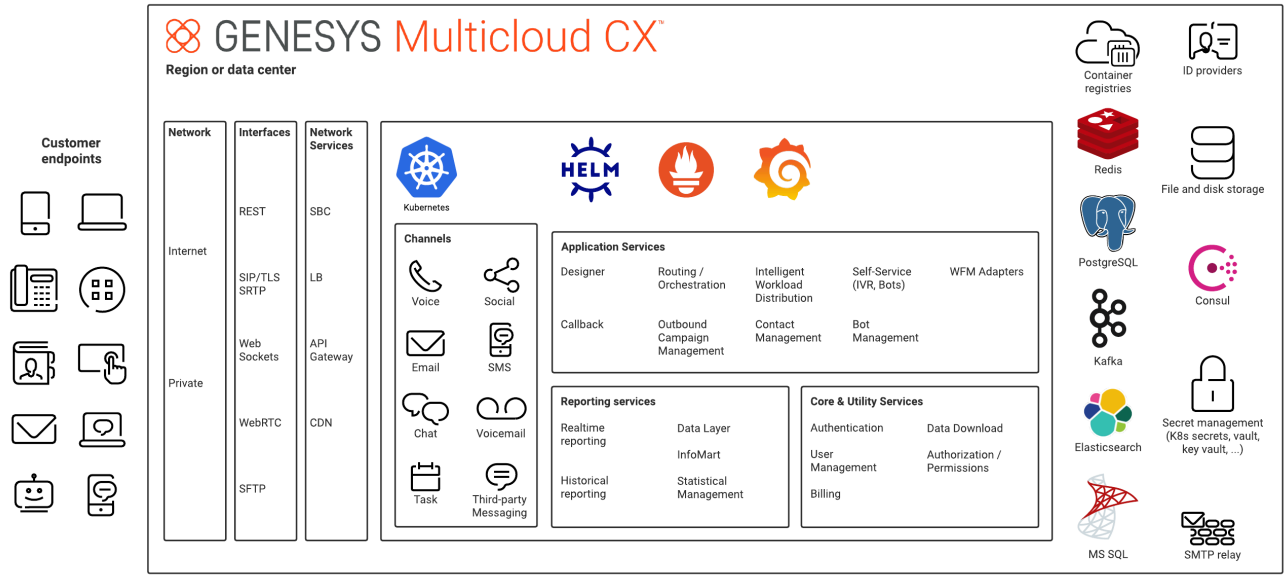
Related documentation:

-
-

RSS:

- For private edition

As mentioned in the About page, Genesys Multicloud CX private edition gives you the flexibility to deploy your contact center on a public cloud or a private one—and even on bare metal servers that reside within your corporate data center.

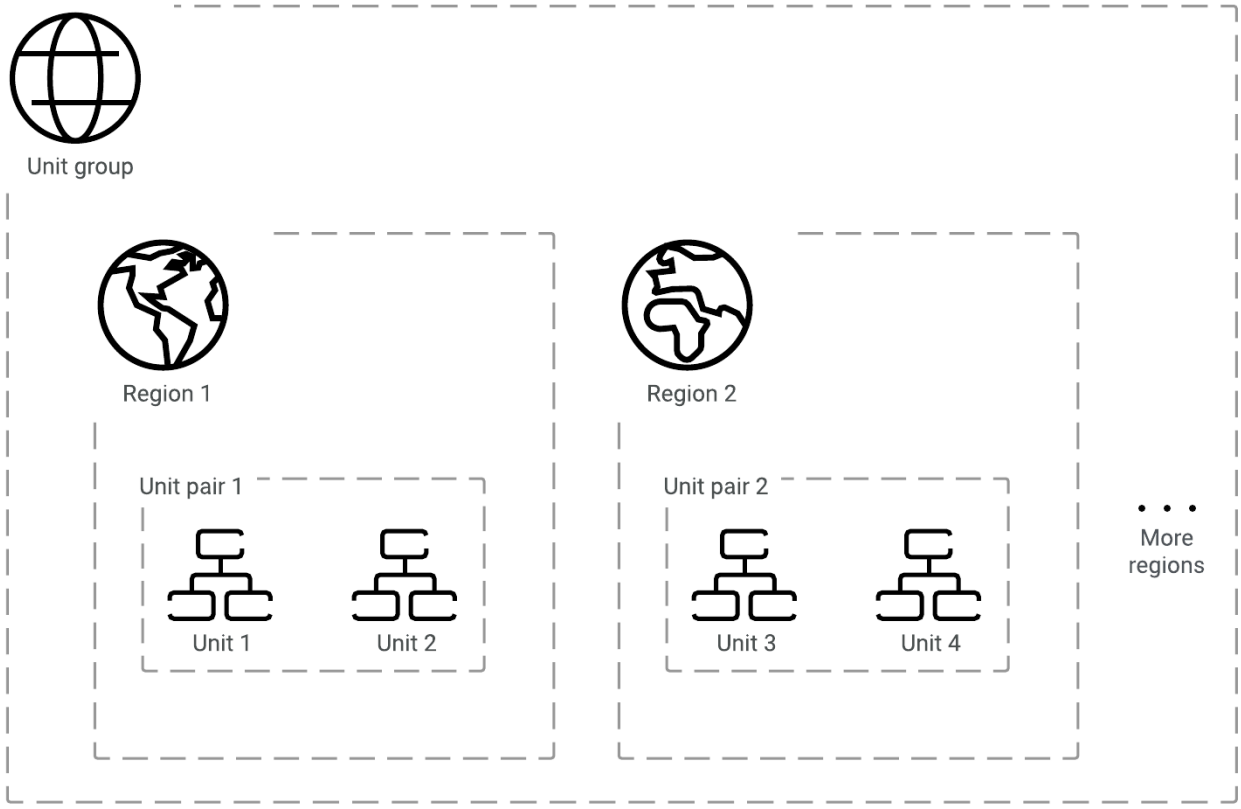


Platform and network

Platform

The basic architecture for private edition involves three levels:

- A **unit** consists of all of the Genesys Multicloud CX and third-party services and resources required to create a single instance of Genesys Multicloud CX private edition. This instance is hosted within a single region or data center.
- A **unit group** brings together a network of units to create a global platform for tenants that covers all geographical regions
- A **unit pair** consists of two units that are part of a unit group and that are both located within a specific geographical region



The following definitions describe important features of the private edition architecture:

- **Region**—A set of isolated and physically separated Availability Zones deployed within a latency-defined perimeter and connected through a dedicated low-latency network within a specific geographical area. **Note:** Regions as defined here are a feature of the cloud deployment architecture and are not supported in the private data center deployment architecture, which does not use Availability Zones.
- **Data center**—A building, a dedicated space within a building, or a group of buildings used to house computer systems and associated components, such as telecommunications and storage systems.
- **Availability Zone (AZ)**—A discrete location within a region that is designed to operate independently from the other Availability Zones in that region. Because of this separation, any given Availability Zone is unlikely to be affected by failures in other Availability Zones. **Note:** Availability Zones are a feature of the cloud deployment architecture and are not supported in the private data center deployment architecture.
  Note: Google Kubernetes Engine (GKE) uses the term "Zone" instead of "Availability Zone."

- **Tenant**—A business entity that has common goals and procedures, and occupies part or all of a contact center. Tenants that share a contact center could be different businesses, or different divisions within the same business.
- **Multi-tenancy**—The partitioning capacity for a platform to host and manage tenants. Each tenant is configured individually and separately.

The following sections provide a more in-depth description of the characteristics of the three levels of the private edition architecture.

Units

A unit can either be dedicated to a specific tenant or used for multiple tenants. The unit and its services and resources can be distributed across Availability Zones if the environment has them.

A unit is composed of the following:

- Network access services (load balancers, firewalls, SBC, and so on)
- A Kubernetes cluster with all of the private edition service pods
- Third-party services (Postgres, Redis, Consul, Kafka, and so on)

There are two main types of units:

- A **primary unit** centralizes certain services used by all regions for a specific tenant, such as Designer application creation, historical reporting, or UI. There is only one primary unit in a unit group. In the current architecture, digital channels are only supported by the primary unit.
- A **secondary unit** only supports voice-related services at this time. Digital channels are only supported by the primary unit.

Unit pairs

Unit pairs provide the following capabilities:

- Redundancy within a geographical region. This geo-redundancy is built into the private edition services.
- Tenants can be distributed across the two units to help reduce the blast area in case of a major failure

A unit pair can consist of a primary unit and a secondary unit, or of two secondary units. **Note:** Unit pairs are only supported by voice-related services at this time.

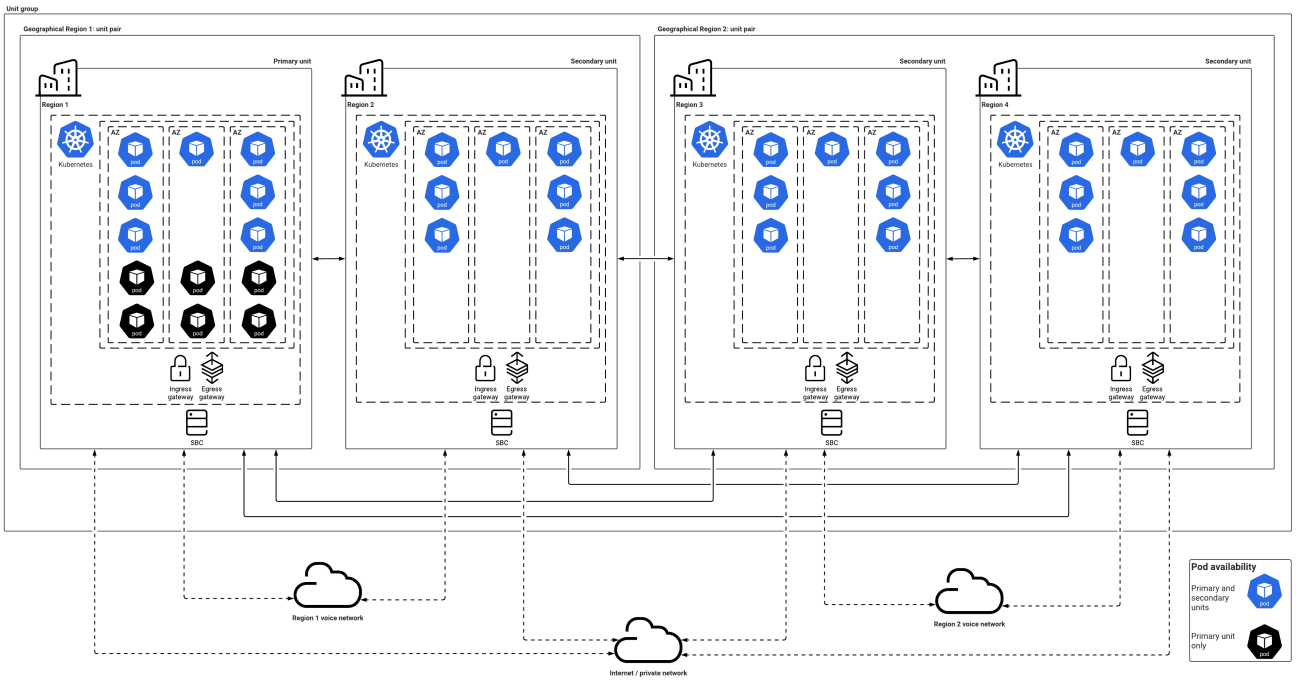
Unit groups

Unit groups interconnect their constituent units by means of a network peering solution, and all inter-region traffic uses either your network connectivity or the network connectivity of your cloud provider. Each group contains a primary unit in one region in the group. This primary regional unit hosts all of the private edition services, while the secondary regional unit hosts only a subset of private edition services. A unit group must contain at least one unit pair. If you add a new geographical region, then you must add a unit pair to the unit group in that geographical region.

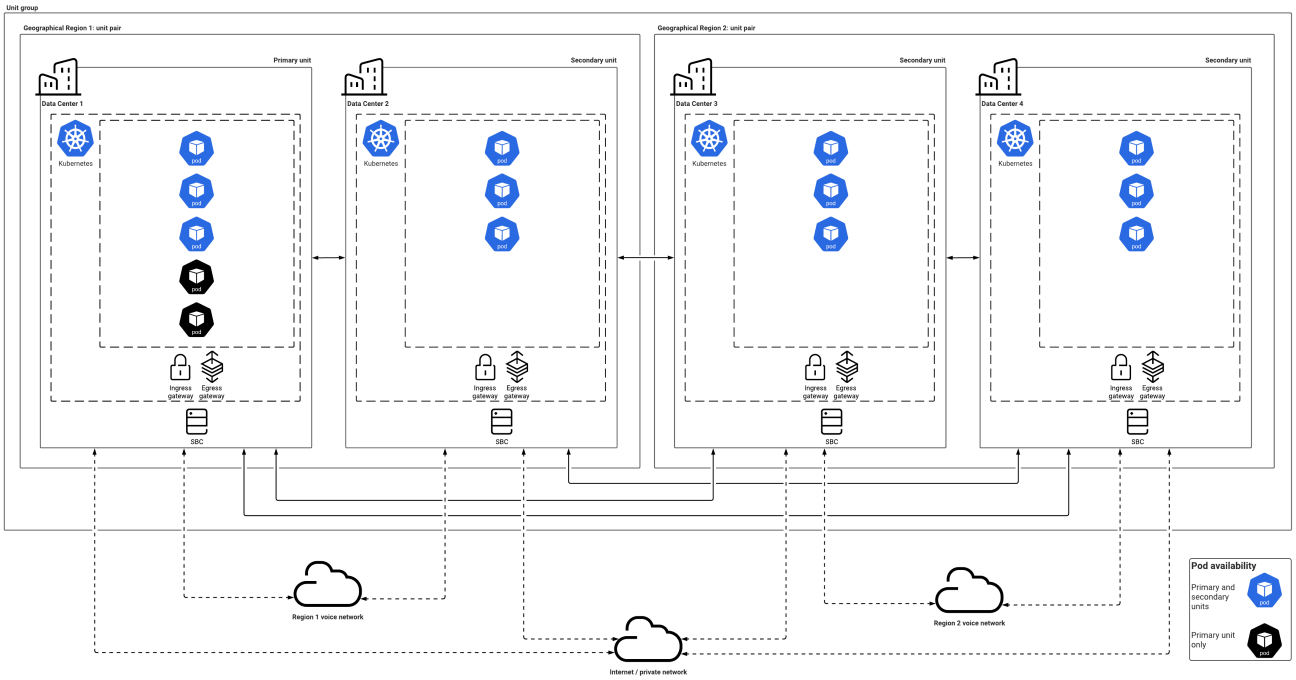
Deployment models

Genesys Multicloud CX private edition allows you to set up a highly available and resilient infrastructure whether you are using a cloud deployment or hosting it in a private data center, as shown in the following diagrams.

Cloud architecture



Private data center architecture



Multiple regions and data centers

The platform supports deployment across multiple regions and data centers. This capability provides extra availability for the voice-related services, with a global view.

- **Call routing and processing**—The ability to distribute call processing across regions. Also, to centrally create and distribute Designer applications across regions.
- **Agent availability**—The ability to have a call processed by agents from any region
- **Data sovereignty**—The ability to contain the data (recordings, and so on) and processing of the call within the region in which the call originated
- **Reporting (Real-time and Historical)**—The ability to provide a global view across all regions
- **Tenant provisioning**—The ability to centrally provision the contact center across multiple regions
- **Callback**—The ability to use a central service to provide in-queue callback across regions

Subnets

Subnets are your responsibility: you must create a subnet for the Kubernetes cluster to accommodate the Genesys Multicloud CX services.

Network access

For information about network access, see Networking Overview.

Supported services

Genesys Multicloud CX private edition supports the services listed on the Genesys Multicloud CX services list.

Software requirements

Genesys Multicloud CX private edition requires the software and versions listed on the software requirements page. Note that you are responsible for installing and deploying the appropriate third-party software in a way that best suits your requirements and the requirements of the Genesys Multicloud CX services.

Kubernetes clusters

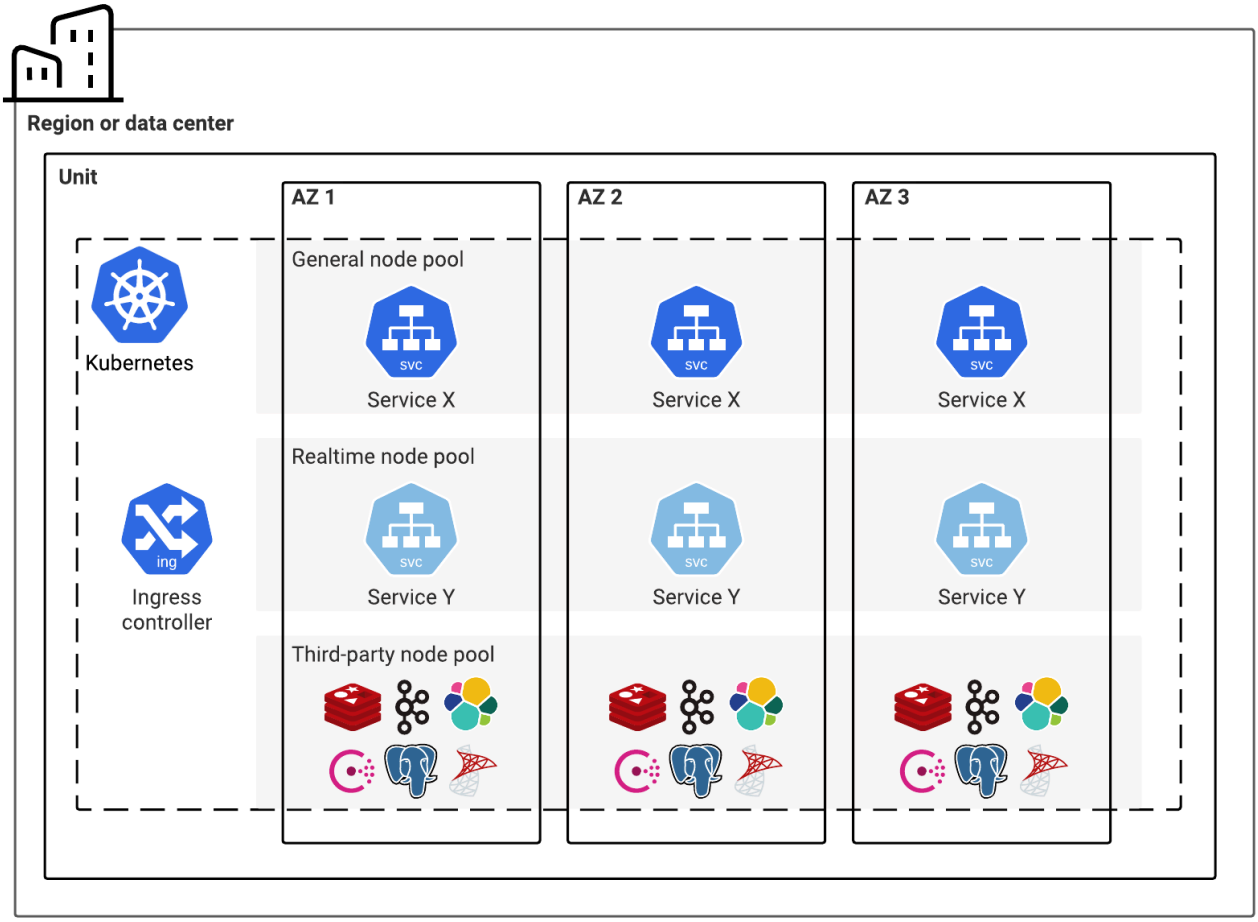
All Genesys Multicloud CX services must run in Kubernetes. Required third-party services can be managed either outside Kubernetes or within Kubernetes. Kubernetes is responsible for managing the running of services, such as monitoring them, restarting them, and so on.

Private edition does not currently support multiple instances of the platform in a single Kubernetes cluster. In other words, if you want to set up separate environments for testing, staging, production, and so on, you must deploy the private edition instances for the various environments in separate clusters.

Deployment

Genesys currently recommends that you use node pools to deploy Kubernetes for the Genesys Multicloud CX services that are hosted within each unit.

- **Node pools**—Genesys recommends that you use the following node pools. Our Helm charts include overrides for the nodeSelector attribute. Use these overrides to assign a service to the appropriate node pool.
  - **General node pool**—This node pool is where most of the Genesys Multicloud CX services are deployed. This type of pool uses general-purpose compute instances with Premium SSD, which provide a 4:1 ratio of memory GB to vCPU.
  - **Real-time node pool**—This node pool is for stateful voice services that require a drain time of 60 minutes or longer to maintain active voice sessions. It uses general-purpose nodes with Premium SSD.
  - **Third-party service node pool (optional)**—This node pool is only needed if you are going to deploy data stores and other third-party services in Kubernetes, such as Redis, Kafka, Postgres or Elasticsearch. These services generally need locally optimized storage and will use the storage-optimized nodes with directly attached NVMe and Premium SSD, which provide an 8:1 ratio of memory GB to vCPU.



Networking

- For information about private edition's general networking requirements and constraints, see Networking Overview
- For information about networking settings for Kubernetes clusters, see Network Settings

Service priorities

For more information, see Service Priorities.

Autoscaling

Most services scale their Kubernetes pods by using the Horizontal Pod Autoscaler. However, this tool can only use CPU or memory metrics from the Kubernetes Metric Server in the HorizontalPodAutoscaler Object. Private edition also works with the Kubernetes cluster scaler. Note that each service provides its own autoscaling rule, and that the autoscaling rule for a specific service is stored in the Helm charts for that service.

Genesys uses the third-party KEDA open-source autoscaler for Genesys Multicloud CX services that require custom metrics from Prometheus. Use the included Helm override attributes to adjust the defaults for each service.

You must perform your own scaling operations on the Kubernetes control plane. The operational requirements of this scaling depend on the size of your contact center. For large installations, you might need to deploy multiple clusters and distribute the Genesys Multicloud CX services across them.

ConfigMaps

Private edition uses ConfigMaps to pass variables and data to the deployed services. This allows each service to be separate from its configuration data, which is a factor in making each service immutable. Genesys provides Helm override attributes that you use to set the configuration values for each service. For more information, see the appropriate service guide.

Operators

You can use operators to deploy most third-party services into clusters. Note that Genesys does not provide operators to deploy Genesys Multicloud CX services.

GKE

> **Important**
> The Genesys implementation of the Google Kubernetes Engine (GKE) is only available on the Google Cloud Platform using the Cloud deployment model.

Notes on what is supported

- Genesys supports deployments to GKE on the Google Cloud Platform (GCP) public cloud
- Genesys recommends using the Container-Optimized OS with ContainerD for Linux nodes.
- Private edition has only been tested on public GKE clusters (public IP addresses for the control plane, worker nodes, and public load balancers) and only supports VPC-native clusters that use alias IP ranges to provide VPC-routable IPs for direct pod access, along with direct access to other Google cloud services. More specifically, private edition does not support routes-based clusters.
- In order to support VPC peering when creating the VPC-native GKE cluster, you must enable IP Alias by including the **--enable-ip-alias** flag, as mentioned in the VPC Peering Restrictions documentation. This eliminates the necessity of creating and exporting routes. By default, VPC network peering with GKE is supported when used with IP aliases.

Not supported

- Although GKE supports a hybrid model, with workloads running both on-premises and in the cloud, private edition does not currently recommend and has not validated splitting its services between different environments or across multiple clusters.

All private edition components must run in a single GKE cluster on the public Google Cloud and not in any private or government GKE instances, or in any GKE instances that are hosted on-premises.

- Private edition does not currently recommend and has not validated hybrid models, or on-premises deployment of GKE.
- GKE clusters in Autopilot mode are not supported.
- GKE Sandbox (gVisor virtualized kernel) and Customer-Managed Keys (CMKs) are not supported.
- Network policies are not provided or supported and all ingress and egress pod traffic must be allowed between all namespaces.
- The GKE Service Mesh add-on, which uses Istio, is not supported. Customers must deploy the Consul Service Mesh instead.
- GKE Ingress is not supported for the initial private edition GKE offering. Instead, private edition requires deploying ingress-nginx.
- Private edition does not support routes-based clusters.

Security

Genesys Multicloud CX private edition has been developed using industry-standard tools and best practices to identify and eliminate security vulnerabilities.

You are responsible for setting up security in the cluster.

For more information about security-related topics, see Security overview.

High-Availability

For more information, see High Availability and Disaster Recovery.

Data stores

Each service must have its own data store cluster or instances, which must not be shared in production environments unless they are under the same service group.

- All data stores must enable and deploy their high availability (HA) functionality
- All data stores must be distributed across Availability Zones, if they are available
- All data stores must support TLS connections and authentication, as appropriate

Here are the data stores used by each service:

Elasticsearch / OpenSearch

> **Important**
> Private edition does not currently support authentication for Elasticsearch or OpenSearch.

The Elasticsearch and OpenSearch services are shareable across tenants, but the tenant data is never shared.

Service	Type of Data	Cross region replication
Designer	Application Analytics data	No
IWD	Interaction and Queue Analytics data	No
TLM	Searchable telemetry data	No
UCS	Searchable contact and interaction history data	No
GWS	Searchable Statistics data	No
CXC	Campaign Analytics	No

Redis

The Redis service is shareable across tenants, but the tenant data is never shared.

Private edition requires the following features for Redis:

- Must support cluster mode
- TLS provisioning
- If you want secure connections to Redis, you must provision Access Control Lists (ACLs) for authentication
- A minimum of three nodes
- A minimum of one replica
- Memory size setting must be based on the services algorithm
- A minimum of two shards per DB
- Must support persistence for services that require it

Service	Type of Data	Cross region access
Pulse	runtime statistics	No
Tenant	stream of tenant data	Yes
CXC	runtime campaign and calling list status	No
Designer	config data	No
GES	runtime callback status and data	No
Nexus	runtime messaging session data	No
IWD	Historical reporting data	No
VMS (all of these services have separate keys (registrate, ORS, ORS stream, Callthread, Agent, Config, SIP, RQ))	runtime interaction, agent, registrations, config and routing request streams, scxml session data	Yes (not all)
GAuth	authentication session data	No
GWS	cached statistics, interaction and agent data	No

SQL databases

> **Important**
>
> All SQL databases except GVP must use Postgres. GVP only supports the use of MS SQL.

You can set up your private edition SQL database instances in either of the following two ways. You can also use the first scenario for some services, and the second scenario for other services:

- Use a separate SQL database instance for each service
- Use a single SQL database instance for a combination of services

However, in each of these scenarios:

- Each service creates its own databases
- Tenant data is never shared.

Service	Type of Data	Shared across tenants	Cross region replication
GCXI	metadata for reports	Yes	No
GVP RS - MS SQL	GVP reporting data	Yes	No
GVP CFG	config data	Yes	No
IXN	digital interaction data	No	No
Pulse Permissions	config data	No	No
Tenant	config and campaign data	No	Yes
GES	config data	Yes	No
GIM	Historical reporting data	No	No
IWD	IWD config data	No	No
Nexus	config data	Yes	No
UCS	config data	Yes	No
UCS	contact, transcriptions, emails, interaction history	No	No
Gauth	config data	Yes	Yes
GWS	config data	Yes	Yes

File and disk storage

For more information, see Storage Requirements.

Voice Connectivity

For more information, see Voice Connectivity.

Email

The following private edition services send emails as part of their service:

- Voicemail
- GCXI
- Pulse

These services use standard mail agents on the operating system over SMTP via ports 25 and 587.

To use email with a service, you must set up the appropriate SMTP relay to relay messages from that service to your email system or email service. **Note:** This must be done from the Kubernetes clusters.

Content delivery networks (CDNs)

The WWE service that runs within private edition delivers static content. You can host this content from a CDN or from NGINX running in the Kubernetes cluster.

Monitoring

Private edition provides appropriate interfaces for you to use your own monitoring tools. For the purposes of this software, monitoring encompasses:

- Metrics
- Logging
- Warnings
- Alerts

Monitoring (metrics)

Private edition provides a set of Prometheus-based metrics and defines an endpoint which the Prometheus platform can scrape. However, it does not provide a Grafana dashboard or Alert rule definitions.

GKE

Private edition uses the Google Cloud operations suite for GKE for system and workload monitoring. The Google Cloud Operations Suite also provides a GKE dashboard for metrics and alerts.

You can enable the GKE workload metrics in order to scrape application metrics based on the PodMonitor resource definition. If a service doesn't provide a PodMonitor resource, then you might need to deploy a Prometheus server with a Stackdriver collector in order to expose Genesys custom application metrics as external metrics in Cloud Monitoring, which does incur an additional cost.

AKS

You can use the Container insights feature in Azure Monitor for monitoring system and workloads in private edition.

Azure Monitor Metrics feature supports collecting metrics from monitored workloads and you can create alerts based on the collected metrics.

Azure Monitor Metrics also supports Prometheus metrics collected from Kubernetes clusters. For more

information, refer Azure product documentation.

Logging

Private edition provides the vast majority of its log data via stdout and stderr. In some exceptional cases, data is logged to disk.

GKE

Private edition uses the Google Cloud operations suite for GKE for system and workload logging. The Google Cloud Operations Suite also provides a Logs Explorer for system and workload logs.

You can either:

- Send your logs to Stdout to be collected and exposed in the Logs Explorer as part of Cloud Logging
- Send them to an RWX/NFS-style log volume provided by a shared Cloud Filestore for legacy or high volume logging.
  Note: RWX/NFS logging will be deprecated in the near future.

AKS

You can use the Log Analytics workspace feature in Azure Monitor for collecting log data of system and workloads in private edition. You can create single or multiple log analytics workspaces based on your organizational needs.

For more information on configuring logs in Azure Monitor log workspaces, refer Azure product documentation.

Integrations

Private edition support integrations with a wide variety of systems to provide an enriched customer experience, including in the following areas:

- Bot platforms, such as Google Dialogflow and AWS Lex
- WFM platforms, such as Verint and Nice
- Email systems
- Identity providers
- Reporting platforms, including business intelligence tools
- Messaging and social platforms
- CRM and BPM systems
- Biometrics systems